

PATENT COOPERATION TREATY

PCT


INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

(Chapter II of the Patent Cooperation Treaty)

(PCT Article 36 and Rule 70)

REC'D 15 DEC 2005

WIP5 PCT

Applicant's or agent's file reference FHG102WO		FOR FURTHER ACTION		See Form PCT/PEA/416
International application No. PCT/EP2004/009221		International filing date (day/month/year) 17.08.2004		Priority date (day/month/year) 19.08.2003
International Patent Classification (IPC) or national classification and IPC G06K9/62				
Applicant FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER .. ET AL				
<p>1. This report is the international preliminary examination report, established by this International Preliminary Examining Authority under Article 35 and transmitted to the applicant according to Article 36.</p> <p>2. This REPORT consists of a total of 8 sheets, including this cover sheet.</p> <p>3. This report is also accompanied by ANNEXES, comprising:</p> <p>a. <input checked="" type="checkbox"/> sent to the applicant and to the International Bureau a total of 6 sheets, as follows:</p> <p><input checked="" type="checkbox"/> sheets of the description, claims and/or drawings which have been amended and are the basis of this report and/or sheets containing rectifications authorized by this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions).</p> <p><input type="checkbox"/> sheets which supersede earlier sheets, but which this Authority considers contain an amendment that goes beyond the disclosure in the international application as filed, as indicated in item 4 of Box No. I and the Supplemental Box.</p> <p>b. <input type="checkbox"/> (sent to the International Bureau only) a total of (indicate type and number of electronic carrier(s)) , containing a sequence listing and/or tables related thereto, in computer readable form only, as indicated in the Supplemental Box Relating to Sequence Listing (see Section 802 of the Administrative Instructions).</p>				
<p>4. This report contains indications relating to the following items:</p> <p><input checked="" type="checkbox"/> Box No. I Basis of the opinion</p> <p><input type="checkbox"/> Box No. II Priority</p> <p><input type="checkbox"/> Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability</p> <p><input type="checkbox"/> Box No. IV Lack of unity of invention</p> <p><input checked="" type="checkbox"/> Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement</p> <p><input type="checkbox"/> Box No. VI Certain documents cited</p> <p><input type="checkbox"/> Box No. VII Certain defects in the international application</p> <p><input type="checkbox"/> Box No. VIII Certain observations on the international application</p>				
Date of submission of the demand 20.06.2005		Date of completion of this report 16.12.2005		
Name and mailing address of the international preliminary examining authority:  European Patent Office - P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Pays Bas Tel. +31 70 340 - 2040 Tx: 31 651 epo nl Fax: +31 70 340 - 3016		Authorized Officer Deltorn, J-M Telephone No. +31 70 340-3468		



**INTERNATIONAL PRELIMINARY REPORT
ON PATENTABILITY**

International application No.
PCT/EP2004/009221

Box No. I Basis of the report

1. With regard to the **language**, this report is based on the international application in the language in which it was filed, unless otherwise indicated under this item.
- ☐ This report is based on translations from the original language into the following language , which is the language of a translation furnished for the purposes of:
- ☐ international search (under Rules 12.3 and 23.1(b))
 - ☐ publication of the international application (under Rule 12.4)
 - ☐ international preliminary examination (under Rules 55.2 and/or 55.3)
2. With regard to the **elements*** of the international application, this report is based on *(replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report):*

Description, Pages

1-52 as originally filed

Claims, Numbers

1-21 received on 20.06.2005 with letter of 20.06.2005

Drawings, Sheets

1-18 as originally filed

- ☐ a sequence listing and/or any related table(s) - see Supplemental Box Relating to Sequence Listing
3. ☐ The amendments have resulted in the cancellation of:
- ☐ the description, pages
 - ☐ the claims, Nos.
 - ☐ the drawings, sheets/figs
 - ☐ the sequence listing (*specify*):
 - ☐ any table(s) related to sequence listing (*specify*):
4. ☐ This report has been established as if (some of) the amendments annexed to this report and listed below had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).
- ☐ the description, pages
 - ☐ the claims, Nos.
 - ☐ the drawings, sheets/figs
 - ☐ the sequence listing (*specify*):
 - ☐ any table(s) related to sequence listing (*specify*):

* If item 4 applies, some or all of these sheets may be marked "superseded."

**INTERNATIONAL PRELIMINARY REPORT
ON PATENTABILITY**

International application No.
PCT/EP2004/009221

Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes: Claims	1-21
	No: Claims	
Inventive step (IS)	Yes: Claims	
	No: Claims	1-21
Industrial applicability (IA)	Yes: Claims	1-21
	No: Claims	

2. Citations and explanations (Rule 70.7):

see separate sheet

Re Item V.

- 1.** The following documents are referred to in this communication:

D1: CAUWENBERGHS G. AND POGGIO T., "Incremental and Decremental Support Vector Machines", ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS , vol. 13, 2001

D2: TAX D M J; DUIN R P.W, "Support vector domain description", PATTERN RECOGNITION LETTERS, vol. 20, nb 11-13, pp. 1191-1199, November 1999

2. CLARITY

The application does not meet the requirements of Article 6 PCT, because claims 1, 13, 15, 19-21 are not clear.

2.1 Claims 1 and 21

The wording "especially the construction of a hypersurface enclosing a finite number of normal objects" of Claim 1, the wording "especially a visual representation" of claims 1 and 21 do not define the matter for which protection is sought.

Furthermore, the wording "predefined optimality condition" used in claims 1 and 21 vague and leaves the reader in doubt as to the meaning of the technical feature to which it refers, thereby rendering the definition of the subject-matter of said claims unclear.

As a consequence, claims 1 and 21 do not meet the requirements of Article 6 PCT.

2.2 Claim 13

The wording "especially sniffing attacks and/or denial of service attacks" of Claim 13

does not define the matter for which protection is sought, hence this claim does not meet the requirements of Article 6 PCT.

2.2 Claim 15

The wording in parenthesis of Claim 15 also does not define the matter for which protection is sought and renders said claim un-concise (Article 6 PCT).

2.3 Claims 19 and 20

The wording in parenthesis ("norm expansion") of claims 19 and 20 renders the subject matter of said claims vague (Article 6 PCT).

3. CLAIMS 1 AND 21

3.1 Claim 1

Furthermore, the above-mentioned lack of clarity notwithstanding, the present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of Claim 1 does not involve an inventive step in the sense of Article 33(3) PCT.

Document D1, which is considered as being the closest prior art to the subject-matter of Claim 1, discloses (the references in parentheses applying to this document):

A method for automatic online detection and classification of objects (abstract) characterised in that:

- (a) the detection of at least one incoming data stream (the concept of a data stream is considered implicit in the sequential training "one vector at a time" of the classifier (see D1, abstract)),
- (b) the automatic construction of a geometric representation of the classification of

the incoming objects of the data stream at a time t_1 subject to at least one predefined optimality condition (section 2.5),

- (d) the online adaptation of the geometric representation of a classification boundary in respect to at least one received object at a time $t_2 > t_1$, the adaptation being subject to at least one predefined optimality condition (section 2.5),
- (e) the online determination of a classification for received objects at t_2 in respect to the geometric representation of normality,
- (f) the automatic classification of normal objects and anomalous objects based on the generated normality classification and generating a data set describing the anomalous data for further processing.

The method of document D1 differs from the method of Claim 1 in that:

- (i) The classification is applied to the detection of outliers in a data stream, whereby the method constructs a geometric representation of normality online.
- (ii) A step (c) consisting in that "the optimal geometric representation of normality is maintained from an instance t_1 after which the construction of the geometric representation of normality is feasible subject to the optimality condition" is introduced between said steps (b) and (d).

The application of the previous classification method to the online one-class classification (i.e., in which a determination of normality is derived from the incoming data only) of anomalous objects (i.e. outliers) in a data stream is disclosed in document D2 (section 2). Determining the conditions (e.g. the minimum number of objects) under which the optimal geometric representation can be found is also disclosed in document D2 (section 3).

The skilled person would therefore regard it as a normal option to include these

features in the method described in document D1 in order to solve the problem of online classification of anomalous objects in a data stream.

3.2 Claim 21

The same reasoning applies to the subject-matter of Claim 21 which corresponds in terms of system to the method of Claim 1. Claim 21, therefore, does not involve an inventive step in the sense of Article 33(3) PCT.

4. CLAIMS 2-20

Dependent claims 2-20 do not contain any features which, in combination with the features of any claim to which they refer, meet the requirements of the PCT in respect of inventive step, the reasons being as follows:

4.1 Claim 2

The feature of Claim 2 wherein the geometric representation of normality is a parametric boundary hypersurface using the enclosure of the minimal volume estimate among all possible surfaces as an optimality condition is also described in document D2 (page 233, line 4 - page 234, line 4; section 2). The subject-matter of Claim 2, therefore, does not involve an inventive step in the sense of Article 33(3) PCT.

4.2 Claims 3-5, 10, 14-16

Document D2 (section 2) also discloses the features of dependent claims 3-5, 10, 14-16. The subject-matter of claims 3-5, 10, 14-16, therefore, does not involve an inventive step in the sense of Article 33(3) PCT.

4.3 Claims 6-9

Document D1 (section 2) discloses the features of dependent claims 6-9. The subject-matter of claims 6-9, therefore, does not involve an inventive step in the sense of Article 33(3) PCT.

4.4 Claims 11-13

The application of the anomaly detection method of Claim 1 to the problem of classifying the logging process (Claim 12), the problem of detecting denial of service attacks (Claim 13), where the data stream of Claim 1 comprises packets in a communication network (Claim 11) cannot be regarded as inventive since it does not presents unexpected effects or properties with respect to the method of Claim 1. The subject-matter of claims 11-13, therefore, does not involve an inventive step in the sense of Article 33(3) PCT.

4.5 Claims 17-20

The features of claims 17-20 consisting in updating the coordinate system (Claim 17), updating the center of coordinates (Claim 18), updating the norms of all objects (Claims 19 and 20) is a matter of normal option that the skilled person would chose when faced with the problem of implementing the online classifier of Claim 14. The subject-matter of claims 17-20, therefore, does not involve an inventive step in the sense of Article 33(3) PCT.

20. 06. 2005

Claims

(76)

1. Method for automatic online detection and classification
of anomalous objects in a data stream, especially comprising
5 datasets and / or signals,
- characterized in
- a) the detection of at least one incoming data stream (1000)
10 containing normal and anomalous objects,
- b) the automatic construction (2100) of a geometric
representation of normality (2200) of the incoming objects of
the data stream (1000) at a time t_1 subject to at least one
15 predefined optimality condition, especially the construction
of a hypersurface enclosing a finite number of normal
objects,
- c) the optimal geometric representation of normality,
20 especially the smallest volume geometric representation of
normality (2200) is maintained from an instance t_i after
which the construction of the geometric representation of
normality (2200) is feasible subject to the optimality
condition,
- 25 d) the online adaptation of the geometric representation of
normality (2200) in respect to at least one received object
at a time $t_2 > t_1$, the adaptation being subject to at least
one predefined optimality condition,
- 30 e) the online determination of a normality/anomaly
classification (2300) for received objects at t_2 in respect
to the geometric representation of normality (2200),
- 35 f) the automatic classification of normal objects and
anomalous objects based on the generated normality
classification (2300) and generating a data set describing

the anomalous data for further processing, especially a visual representation.

2. Method according to claim 1, characterised in that
5 the geometric representation of normality (2200) is a parametric boundary hypersurface using the enclosure of the minimal volume or the minimal volume estimate among all possible surfaces as an optimality condition.
- 10 3. Method according to claim 2, characterised in that the hypersurface is constructed in the space of original measurements of least one incoming data stream (1000) or in a space obtained by a nonlinear transformation thereof.
- 15 4. Method according to at least one preceding claim, characterised in that the optimality condition, used to construct the parametric boundary hypersurface, is a pre-defined condition, especially the one based on an expected fraction η of anomalous objects, or a condition, dynamically
20 adaptable to the data stream.
5. Method according to at least one preceding claim, characterised in that the anomalous objects are determined as the ones lying outside of the geometrical
25 representation of normality (2200), especially the parametric boundary hypersurface enclosing the normal objects.
6. Method according to at least one preceding claim, characterized in that dynamic adaptation of the
30 geometric representation of normality (2200) comprises an automatic adjustment of parameters x_i of the geometric representation of normality (2200) to incorporate at least one new object while maintaining the optimality of the geometric representation of normality (2200).
35
7. Method according to at least one preceding claim, characterized in that the dynamic adaptation of the

geometric representation of normality (2200) comprises an automatic adjustment of parameters x_i of the geometric representation of normality (2200) to remove the least-relevant object while maintaining the optimality of the
5 geometric representation of normality (2200).

8. Method according to at least one preceding claim, characterized in that the geometric representation of normality (2200) is generated with a Support Vector Machine
10 method, generating a parametric vector x to describe the representation.

9. Method according to at least one preceding claim, characterised in that the temporal change of the
15 geometrical representation of normality (2200), especially the temporal change of a parameter vector x of the geometrical representation of normality (2200) is stored for the evaluation of temporal trend in the data stream (1000).

20 10. Method according to at least on one preceding claim, characterised in that the geometric representation of normality (2200) is a sphere or any part thereof.

11. Method according to at least one preceding claim,
25 characterized in that incoming data stream (1000) comprises data packets in communication networks or representations thereof.

12. Method according to at least one preceding claim,
30 characterized in that the data objects comprises entries originating from the logging in process in at least one computer or representations thereof.

13. Method according to claim 11 or 12, characterized in
35 that the determination of normality of the received data packets distinguishes normal incoming data stream from anomalous data, especially sniffing attacks and / or denial

of service attacks, whereby the means for automatic determining the normal and anomalous data generates a warning message.

5 14. A method according to any preceding claim,
characterized in that, the method for construction and
update of the geometric representation of normality (2200) in
which the coordinate system in which the representation is
constructed is fixed to some point in the data space or in
10 the feature space.

15 15. A method according to claim 14, in which the center of
coordinate system coincides with the center of mass of the
data space (in the original or in the feature space)

16. A method according to claim 14 or 15, in which the
decision on normality or anomaly of an object is decided
upon its norm in the data-centered (or feature-space-
centered) coordinate system, or by the radius of the
20 hypersphere centered at the center of the origin in the said
coordinate system and encompassing the given objects.

25 17. A method according to one of the claims 14 to 16 in which
the update of the representation includes the update of the
coordinate system.

30 18. A method according to one of the claims 14 to 17 in which
the update of coordinate system includes the update of the
center of coordinates.

35 19. A method according to one of the claims 14 to 18 in which
importation of the new object includes as a part the update
of the norms of all objects in the working set so as to bring
them in the new coordinate system corresponding to the
expanded working set ("norm expansion").

20. A method according to one of the claims 15 to 19, in

which removal of the object includes as a part the update of the norms of all objects in the working set so as to bring them in the new coordinate system corresponding to the contracted working set ("norm contraction")

5

21. System for automatic online detection and classification of anomalous objects in a data stream, especially comprising datasets and / or signals,

10 characterized by

a) a detection means for least one incoming data stream (1000) containing normal and anomalous objects,

15 b) an automatic online anomaly detection engine comprising

20 - an automatic construction means (2100) of a geometric representation of normality (2200) for the incoming objects of the data stream (1000) at a time t_1 subject to at least one predefined optimality condition, especially for the construction of a hypersurface enclosing a finite number of normal objects, with an automatic online adaptation means for the geometric representation of normality (2200) in respect to received at least one
25 received object at a time $t_2 > t_1$, the adaptation being subject to at least one predefined optimality condition, and

30 - the optimal geometric representation of normality, especially the smallest volume geometric representation of normality (2200) is maintained from an instance t_i after which the construction of the geometric representation of normality (2200) is feasible subject to the optimality condition,

35 - an automatic online determination means of a normality classification (2300) for received objects at t_2 in

respect to the geometric representation of normality
(2200).

- 5 c) an automatic classification means (4000) of normal objects
and anomalous objects based on the generated normality
classification (2300) and generating a data set describing
the anomalous data for further processing, especially a
visual representation.